

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/802,472

DATE: 07/09/2001

TIME: 12:44:34

Input Set : A:\EINAT4-1C.txt

Output Set: N:\CRF3\07092001\I802472.raw

Does Not Comply
Corrected Diskette Needed

pg 3, 5-6

3 <110> APPLICANT: EINAT, Paz
 4 SKALITER, Rami
 5 FEINSTEIN, Elena
 7 <120> TITLE OF INVENTION: SEQUENCES CHARACTERISTIC OF HYPOXIA-REGULATED GENE
 TRANSCRIPTION

9 <130> FILE REFERENCE: EINAT=4.1C
 11 <140> CURRENT APPLICATION NUMBER: US/09/802,472
 11 <141> CURRENT FILING DATE: 2001-03-09
 11 <150> PRIOR APPLICATION NUMBER: US 09/383,096
 12 <151> PRIOR FILING DATE: 1999-08-27
 14 <150> PRIOR APPLICATION NUMBER: US 09/138,109
 15 <151> PRIOR FILING DATE: 1998-08-21
 17 <150> PRIOR APPLICATION NUMBER: US 60/098,158
 18 <151> PRIOR FILING DATE: 1998-08-27
 20 <150> PRIOR APPLICATION NUMBER: US 60/132,684
 21 <151> PRIOR FILING DATE: 1999-05-05
 23 <160> NUMBER OF SEQ ID NOS: 21
 25 <170> SOFTWARE: PatentIn version 3.0
 27 <210> SEQ ID NO: 1
 28 <211> LENGTH: 1655
 29 <212> TYPE: DNA
 30 <213> ORGANISM: Homo sapiens
 32 <220> FEATURE:
 33 <221> NAME/KEY: CDS
 34 <222> LOCATION: (265)..(1575)
 36 <220> FEATURE:
 37 <221> NAME/KEY: misc feature
 38 <223> OTHER INFORMATION: "n" at every occurrence is unknown
 41 <400> SEQUENCE: 1
 42 gcacnaggtg tgtggcagca anagccgccg gttcgggacc nccgcactg ggggtggcaac 60
 44 ggcgcaggag gggctgcggg gagggagtg tgagcgcagg cggcaggggt ctgggaaaga 120
 46 cgaagtcgct atttgctgtc tgagcgcgct cgcagctcct ggaagtgttg ccgcctctcg 180
 48 gtttcgctct cgctcgctgc gctcctagaa ggggcggccg cctccaggac tgaccagggc 240
 50 caagtggcgc tcggcgggca ctac atg gcg gag ggt gaa ggg tac ttc gcc 291
 51 Met Ala Glu Gly Glu Gly Tyr Phe Ala
 52 1 5
 54 atg tct gag gac gag ctg gcc tgc agc ccc tac atc ccc cta ggc ggc 339
 55 Met Ser Glu Asp Glu Leu Ala Cys Ser Pro Tyr Ile Pro Leu Gly Gly
 56 10 15 20 25
 58 gac ttc ggc ggc ggc gac ttc ggc ggc ggc gac ttc ggc ggt ggc ggc 387
 59 Asp Phe Gly Gly Gly Asp Phe Gly Gly Gly Asp Phe Gly Gly Gly Gly
 60 30 35 40
 62 agc ttc ggt ggg cat tgc ttg gac tat tgc gaa agc cct acg gcg cac 435
 63 Ser Phe Gly Gly His Cys Leu Asp Tyr Cys Glu Ser Pro Thr Ala His
 64 45 50 55
 66 tgc aat gtg ctg aac tgg gag caa gtg cag cgg ctg gac ggc atc ctg 483
 67 Cys Asn Val Leu Asn Trp Glu Gln Val Gln Arg Leu Asp Gly Ile Leu
 68 60 65 70

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/802,472

DATE: 07/09/2001

TIME: 12:44:34

Input Set : A:\EINAT4-1C.txt

Output Set: N:\CRF3\07092001\I802472.raw

```

70 agc gag acc att ccg att cac ggg cgc ggc aac ttc ccc acg ctc gag      531
71 Ser Glu Thr Ile Pro Ile His Gly Arg Gly Asn Phe Pro Thr Leu Glu
72      75      80      85
74 ctg cag ccg agc ctg atc gtg aag gtg gtg cgg cgg cgc ctg gcc gag      579
75 Leu Gln Pro Ser Leu Ile Val Lys Val Val Arg Arg Arg Leu Ala Glu
76 90      95      100      105
78 aag cgc att ggc gtc cgc gac gtg cgc ctc aac ggc tgc gca gcc agc      627
79 Lys Arg Ile Gly Val Arg Asp Val Arg Leu Asn Gly Ser Ala Ala Ser
80      110      115      120
82 cat gtc ctg cac cag gac agc ggc ctg ggc tac aag gac ctg gac ctc      675
83 His Val Leu His Gln Asp Ser Gly Leu Gly Tyr Lys Asp Leu Asp Leu
84      125      130      135
86 atc ttc tgc gcc gac ctg cgc ggg gaa ggg gag ttt cag act gtg aag      723
87 Ile Phe Cys Ala Asp Leu Arg Gly Glu Gly Glu Phe Gln Thr Val Lys
88      140      145      150
90 gac gtc gtg ctg gac tgc ctg ttg gac ttc tta ccc gag ggg gtg aac      771
91 Asp Val Val Leu Asp Cys Leu Leu Asp Phe Leu Pro Glu Gly Val Asn
92      155      160      165
94 aaa gag aag atc aca cca ctc acg ctc aag gaa gct tat gtg cag aaa      819
95 Lys Glu Lys Ile Thr Pro Leu Thr Leu Lys Glu Ala Tyr Val Gln Lys
96 170      175      180      185
98 atg gtt aaa gtg tgc aat gac tct gac cga tgg agt ctt ata tcc ctg      867
99 Met Val Lys Val Cys Asn Asp Ser Asp Arg Trp Ser Leu Ile Ser Leu
100      190      195      200
102 tca aac aac agt ggc aaa aat gtg gaa ctg aaa ttt gtg gat tcc ctc      915
103 Ser Asn Asn Ser Gly Lys Asn Val Glu Leu Lys Phe Val Asp Ser Leu
104      205      210      215
106 cgg agg cag ttt gaa ttc agt gta gat tct ttt caa atc aaa tta gac      963
107 Arg Arg Gln Phe Glu Phe Ser Val Asp Ser Phe Gln Ile Lys Leu Asp
108      220      225      230
110 tct ctt ctg ctc ttt tat gaa tgt tca gag aac cca atg act gag aca      1011
111 Ser Leu Leu Leu Phe Tyr Glu Cys Ser Glu Asn Pro Met Thr Glu Thr
112      235      240      245
114 ttt cac ccc aca ata atc ggg gag agc gtc tat ggc gat ttc cag gaa      1059
115 Phe His Pro Thr Ile Ile Gly Glu Ser Val Tyr Gly Asp Phe Gln Glu
116 250      255      260      265
118 gcc ttt gat cac ctt tgt aac aag atc att gcc acc agg aac cca gag      1107
119 Ala Phe Asp His Leu Cys Asn Lys Ile Ile Ala Thr Arg Asn Pro Glu
120      270      275      280
122 gaa atc cga ggg gga ggc ctg ctt aag tac tgc aac ctc ttg gtg agg      1155
123 Glu Ile Arg Gly Gly Gly Leu Leu Lys Tyr Cys Asn Leu Leu Val Arg
124      285      290      295
126 ggc ttt agg ccc gcc tct gat gaa atc aag acc ctt caa agg tat atg      1203
127 Gly Phe Arg Pro Ala Ser Asp Glu Ile Lys Thr Leu Gln Arg Tyr Met
128      300      305      310
130 tgt tcc agg ttt ttc atc gac ttc tca gac att gga gag cag cag aga      1251
131 Cys Ser Arg Phe Phe Ile Asp Phe Ser Asp Ile Gly Glu Gln Gln Arg
132      315      320      325
134 aaa ctg gag tcc tat ttg cag aac ctc ttt gtg gga ttg gaa gcc cgc      1299

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/802,472

DATE: 07/09/2001

TIME: 12:44:34

Input Set : A:\EINAT4-1C.txt

Output Set: N:\CRF3\07092001\I802472.raw

```

135 Lys Leu Glu Ser Tyr Leu Gln Asn Leu Phe Val Gly Leu Glu Ala Arg
136 330 335 340 345
138 aag tat gag tat ctc atg acc ctt cat gga gtg gta aat gag agc tca 1347
139 Lys Tyr Glu Tyr Leu Met Thr Leu His Gly Val Val Asn Glu Ser Ser
140 350 355 360
142 gtg tgc ctg atg gga cat gaa aga aga cag act tta aac ctt atc acc 1395
143 Val Cys Leu Met Gly His Glu Arg Arg Gln Thr Leu Asn Leu Ile Thr
144 365 370 375
146 atg ctg gct atc cgg gtg tta gct gac caa aat gtc att cct aat gtg 1443
147 Met Leu Ala Ile Arg Val Leu Ala Asp Gln Asn Val Ile Pro Asn Val
148 380 385 390
150 gct aat gtc act tgc tat tac cag cca gcc ccc tat gta gca gat gcc 1491
151 Ala Asn Val Thr Cys Tyr Tyr Gln Pro Ala Pro Tyr Val Ala Asp Ala
152 395 400 405
154 aac ttt agc aat tac tac att gca cag gtt cag cca gta ttc acg tgc 1539
155 Asn Phe Ser Asn Tyr Tyr Ile Ala Gln Val Gln Pro Val Phe Thr Cys
156 410 415 420 425
158 cag caa cag acc tac tcc act tgg cta ccc tgc aat taagaatcat 1585
159 Gln Gln Gln Thr Tyr Ser Thr Trp Leu Pro Cys Asn
160 430 435
162 ttaaaaatgt cctgtgggga agccatttca gacaagacag gagagaaaaa aaaaaaaaaa 1645
164 aaaaaaaaaa 1655
167 <210> SEQ ID NO: 2
168 <211> LENGTH: 437
169 <212> TYPE: PRT
170 <213> ORGANISM: Homo sapiens
172 <220> FEATURE:
173 <221> NAME/KEY: misc feature
174 <223> OTHER INFORMATION: "n" at every occurrence is unknown
176 <400> SEQUENCE: 2
178 Met Ala Glu Gly Glu Gly Tyr Phe Ala Met Ser Glu Asp Glu Leu Ala
179 1 5 10 15
182 Cys Ser Pro Tyr Ile Pro Leu Gly Gly Asp Phe Gly Gly Gly Asp Phe
183 20 25 30
186 Gly Gly Gly Asp Phe Gly Gly Gly Gly Ser Phe Gly Gly His Cys Leu
187 35 40 45
190 Asp Tyr Cys Glu Ser Pro Thr Ala His Cys Asn Val Leu Asn Trp Glu
191 50 55 60
194 Gln Val Gln Arg Leu Asp Gly Ile Leu Ser Glu Thr Ile Pro Ile His
195 65 70 75 80
198 Gly Arg Gly Asn Phe Pro Thr Leu Glu Leu Gln Pro Ser Leu Ile Val
199 85 90 95
202 Lys Val Val Arg Arg Arg Leu Ala Glu Lys Arg Ile Gly Val Arg Asp
203 100 105 110
206 Val Arg Leu Asn Gly Ser Ala Ala Ser His Val Leu His Gln Asp Ser
207 115 120 125
210 Gly Leu Gly Tyr Lys Asp Leu Asp Leu Ile Phe Cys Ala Asp Leu Arg
211 130 135 140
214 Gly Glu Gly Glu Phe Gln Thr Val Lys Asp Val Val Leu Asp Cys Leu

```

*delete-amino acid sequences
would contain
Xaa, not "n"*

*Besides, no
Xaa is
shown in
the sequence*

*"n" is only
used in
nucleotide
sequence*

RAW SEQUENCE LISTING

DATE: 07/09/2001

PATENT APPLICATION: US/09/802,472

TIME: 12:44:34

Input Set : A:\EINAT4-1C.txt

Output Set: N:\CRF3\07092001\I802472.raw

```

215 145          150          155          160
218 Leu Asp Phe Leu Pro Glu Gly Val Asn Lys Glu Lys Ile Thr Pro Leu
219          165          170          175
222 Thr Leu Lys Glu Ala Tyr Val Gln Lys Met Val Lys Val Cys Asn Asp
223          180          185          190
226 Ser Asp Arg Trp Ser Leu Ile Ser Leu Ser Asn Asn Ser Gly Lys Asn
227          195          200          205
230 Val Glu Leu Lys Phe Val Asp Ser Leu Arg Arg Gln Phe Glu Phe Ser
231          210          215          220
234 Val Asp Ser Phe Gln Ile Lys Leu Asp Ser Leu Leu Leu Phe Tyr Glu
235 225          230          235          240
238 Cys Ser Glu Asn Pro Met Thr Glu Thr Phe His Pro Thr Ile Ile Gly
239          245          250          255
242 Glu Ser Val Tyr Gly Asp Phe Gln Glu Ala Phe Asp His Leu Cys Asn
243          260          265          270
246 Lys Ile Ile Ala Thr Arg Asn Pro Glu Glu Ile Arg Gly Gly Gly Leu
247          275          280          285
250 Leu Lys Tyr Cys Asn Leu Leu Val Arg Gly Phe Arg Pro Ala Ser Asp
251          290          295          300
254 Glu Ile Lys Thr Leu Gln Arg Tyr Met Cys Ser Arg Phe Phe Ile Asp
255 305          310          315          320
258 Phe Ser Asp Ile Gly Glu Gln Gln Arg Lys Leu Glu Ser Tyr Leu Gln
259          325          330          335
262 Asn Leu Phe Val Gly Leu Glu Ala Arg Lys Tyr Glu Tyr Leu Met Thr
263          340          345          350
266 Leu His Gly Val Val Asn Glu Ser Ser Val Cys Leu Met Gly His Glu
267          355          360          365
270 Arg Arg Gln Thr Leu Asn Leu Ile Thr Met Leu Ala Ile Arg Val Leu
271          370          375          380
274 Ala Asp Gln Asn Val Ile Pro Asn Val Ala Asn Val Thr Cys Tyr Tyr
275 385          390          395          400
278 Gln Pro Ala Pro Tyr Val Ala Asp Ala Asn Phe Ser Asn Tyr Tyr Ile
279          405          410          415
282 Ala Gln Val Gln Pro Val Phe Thr Cys Gln Gln Gln Thr Tyr Ser Thr
283          420          425          430
286 Trp Leu Pro Cys Asn
287          435
290 <210> SEQ ID NO: 3
291 <211> LENGTH: 3454
292 <212> TYPE: DNA
293 <213> ORGANISM: Homo sapiens
295 <220> FEATURE:
296 <221> NAME/KEY: CDS
297 <222> LOCATION: (323)..(1762)
299 <220> FEATURE:
300 <221> NAME/KEY: misc_feature
301 <223> OTHER INFORMATION: "n" at every occurrence is unknown
304 <400> SEQUENCE: 3
305 ctccgcggcg gggatgctga ggagcgctgg gtccgggagc agccctggcc cctgcggact

```

60

RAW SEQUENCE LISTING

DATE: 07/09/2001

PATENT APPLICATION: US/09/802,472

TIME: 12:44:34

Input Set : A:\EINAT4-1C.txt

Output Set: N:\CRF3\07092001\I802472.raw

```

307 tccgaggccg tgaaaacccc tgcgctgcgg cccttcccag gcccccgagg ccgttcgccc 120
309 ttcccgaagc ccgactgggg gaagagtcca gcaccaaagc ggccgttctc ggattccgga 180
311 gcgttctgga gccccgagag acgccccggg gttctagaag ctccccggcg gcgcccagtc 240
313 ccggttcat tcgggcgtcc ctccgaaacc cactcgggtg cacgggtcgt cggcgagccc 300
315 cgaccgggtc ctggcgcgca cc atg atc gtg gcg gac tcc gag tgc cgc gca 352
316                               Met Ile Val Ala Asp Ser Glu Cys Arg Ala
317                               1                               5                               10
319 gag ctc aag gac tac ctg cgg ttc gcc ccg ggc ggc gtc ggc gac tcg 400
320 Glu Leu Lys Asp Tyr Leu Arg Phe Ala Pro Gly Gly Val Gly Asp Ser
321                               15                               20                               25
323 ggc ccc gga gag gag cag agg gag agc cgg gct cgg cga ggc cct cga 448
324 Gly Pro Gly Glu Glu Gln Arg Glu Ser Arg Ala Arg Arg Gly Pro Arg
325                               30                               35                               40
327 ggg ccc agc gcc ttc atc ccc gtg gag gag gtc ctt cgg gag ggg gct 496
328 Gly Pro Ser Ala Phe Ile Pro Val Glu Glu Val Leu Arg Glu Gly Ala
329                               45                               50                               55
331 gag agc ctc gag cag cac ctg ggg ctg gag gca ctg atg tcc tct ggg 544
332 Glu Ser Leu Glu Gln His Leu Gly Leu Glu Ala Leu Met Ser Ser Gly
333                               60                               65                               70
335 cga gta gac aac ctg gca gtg gtg atg ggc ctg cac cct gac tac ttt 592
336 Arg Val Asp Asn Leu Ala Val Val Met Gly Leu His Pro Asp Tyr Phe
337 75                               80                               85                               90
339 acc agc ttc tgg cnc ctg cac tac ctg ctg ctg cac acg gat ggt ccc 640
340 Thr Ser Phe Trp Xaa Leu His Tyr Leu Leu Leu His Thr Asp Gly Pro
341                               95                               100                               105
343 ttg gcc agc tcc tgg cgc cac tac att gcc atc atg gct gcc gcc cgc 688
344 Leu Ala Ser Ser Trp Arg His Tyr Ile Ala Ile Met Ala Ala Ala Arg
345                               110                               115                               120
347 cat cag tgt tct tac ctg gta ggc tcc cac atg gcc gag ttt ctg cag 736
348 His Gln Cys Ser Tyr Leu Val Gly Ser His Met Ala Glu Phe Leu Gln
349                               125                               130                               135
351 act ggt ggt gac cct gag tgg ctg ctg ggc ctc cac cgg gcc ccc gag 784
352 Thr Gly Gly Asp Pro Glu Trp Leu Leu Gly Leu His Arg Ala Pro Glu
353                               140                               145                               150
355 aag ctg cgc aaa ctc agc gag atc aac aag ttg ctg gcg cat cgg cca 832
356 Lys Leu Arg Lys Leu Ser Glu Ile Asn Lys Leu Leu Ala His Arg Pro
357 155                               160                               165                               170
359 tgg ctc atc acc aag gaa cac atc cag gcc ttg ctg aag acc ggc gag 880
360 Trp Leu Ile Thr Lys Glu His Ile Gln Ala Leu Leu Lys Thr Gly Glu
361                               175                               180                               185
363 cac act tgg tcc ctg gcc gag ctc att cag gct ctg gtc ctg ctc acc 928
364 His Thr Trp Ser Leu Ala Glu Leu Ile Gln Ala Leu Val Leu Leu Thr
365                               190                               195                               200
367 cac tgc cac tcg ctc tcc tcc ttc gtg ttt ggc tgt ggc atc ctc cct 976
368 His Cys His Ser Leu Ser Ser Phe Val Phe Gly Cys Gly Ile Leu Pro
369                               205                               210                               215
371 gag ggg gat gca gat ggc agc cct gcc ccc cag gca cct aca ccc cct 1024
372 Glu Gly Asp Ala Asp Gly Ser Pro Ala Pro Gln Ala Pro Thr Pro Pro
373                               220                               225                               230

```

Use of n and/or Xaa has been detected in the Sequence Listing.
 Review the Sequence Listing to insure a corresponding
 explanation is presented in the <220> to <223> fields of
 each sequence using n or Xaa.

09/802,472 6

<210> 4
<211> 480
<212> PRT
<213> Homo sapiens

<220>

<221> misc feature

<223> "n" at every occurrence is unknown

This is an amino acid sequence;

<400> 4

Xaa is shown below

Met Ile Val Ala Asp Ser Glu Cys Arg Ala Glu Leu Lys Asp Tyr Leu
1 5 10 15

Arg Phe Ala Pro Gly Gly Val Gly Asp Ser Gly Pro Gly Glu Glu Gln
20 25 30

Arg Glu Ser Arg Ala Arg Arg Gly Pro Arg Gly Pro Ser Ala Phe Ile
35 40 45

Pro Val Glu Glu Val Leu Arg Glu Gly Ala Glu Ser Leu Glu Gln His
50 55 60

Leu Gly Leu Glu Ala Leu Met Ser Ser Gly Arg Val Asp Asn Leu Ala
65 70 75 80

Val Val Met Gly Leu His Pro Asp Tyr Phe Thr Ser Phe Trp Xaa Leu
85 90 95

His Tyr Leu Leu Leu His Thr Asp Gly Pro Leu Ala Ser Ser Trp Arg
100 105 110

His Tyr Ile Ala Ile Met Ala Ala Ala Arg His Gln Cys Ser Tyr Leu
115 120 125

Val Gly Ser His Met Ala Glu Phe Leu Gln Thr Gly Gly Asp Pro Glu
130 135 140

Trp Leu Leu Gly Leu His Arg Ala Pro Glu Lys Leu Arg Lys Leu Ser
145 150 155 160

Glu Ile Asn Lys Leu Leu Ala His Arg Pro Trp Leu Ile Thr Lys Glu
165 170 175

His Ile Gln Ala Leu Leu Lys Thr Gly Glu His Thr Trp Ser Leu Ala
180 185 190

Use of n and/or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to insure a corresponding explanation is presented in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/802,472

DATE: 07/09/2001

TIME: 12:44:35

Input Set : A:\EINAT4-1C.txt

Output Set: N:\CRF3\07092001\I802472.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:42 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:339 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:340 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:463 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:465 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:527 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:4
L:527 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:783 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:6
L:783 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:785 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:6
L:785 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:821 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:6
L:821 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:823 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:6
L:823 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:825 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:6
L:825 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:861 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:7
L:861 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:2527 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:20
L:2539 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:20
L:2539 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20